

## RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.

Application Serial Number: 10/089,057A  
Source: 1FW16  
Date Processed by STIC: 10/18/04

# ***ENTERED***



IFW16

## RAW SEQUENCE LISTING

DATE: 10/18/2004

PATENT APPLICATION: US/10/089,057A

TIME: 09:38:04

Input Set : A:\PTO.FG.txt

Output Set: N:\CRF4\10182004\J089057A.raw

3 <110> APPLICANT: HIRANO, SEIKO  
 4 KIMURA, EIICHIRO  
 5 OSUMI, TSUYOSHI  
 6 MATSUI, KAZUHIKO  
 7 KAWAHARA, YOSHIO  
 8 NONAKA, GEN  
 9 MATSUZAKI, YUMI  
 10 AKIYOSHI, NAOKI  
 11 NAKAMURA, KANAE  
 12 KURAHASHI, OSAMU  
 13 NAKAMATSU, TSUYOSHI  
 14 SUGIMOTO, SHINICHI  
 16 <120> TITLE OF INVENTION: GENES FOR HEAT RESISTANT ENZYMES OF AMINO ACID BIOSYNTHETIC  
 PATHWAY  
 17 DERIVED FROM THERMOPHILIC CORYNEFORM BACTERIA  
 W--> 18 <130> FILE REFERENCE: 221519US0PCT  
 20 <140> CURRENT APPLICATION NUMBER: 10/089,057A  
 21 <141> CURRENT FILING DATE: 2002-04-03  
 23 <150> PRIOR APPLICATION NUMBER: PCT/JP00/06913  
 24 <151> PRIOR FILING DATE: 2000-10-04  
 26 <150> PRIOR APPLICATION NUMBER: JP 11-282716  
 27 <151> PRIOR FILING DATE: 1999-10-04  
 29 <150> PRIOR APPLICATION NUMBER: JP 11-311147  
 30 <151> PRIOR FILING DATE: 1999-11-01  
 32 <150> PRIOR APPLICATION NUMBER: JP 2000-120687  
 33 <151> PRIOR FILING DATE: 2000-04-21  
 35 <160> NUMBER OF SEQ ID NOS: 108  
 37 <170> SOFTWARE: PatentIn version 3.1  
 39 <210> SEQ ID NO: 1  
 40 <211> LENGTH: 1980  
 41 <212> TYPE: DNA  
 42 <213> ORGANISM: Corynebacterium thermoaminogenes  
 44 <220> FEATURE:  
 45 <221> NAME/KEY: CDS  
 46 <222> LOCATION: (577)..(1869)  
 48 <400> SEQUENCE: 1  
 49 tgcattccac cgacgggtcac gcgttcgggtc ttgtcagcgg cgtcaatctg ctgatgggtc 60  
 50 atgcaaagct ccttcgaagc aagagatcgg gtgtgtgcgg gcacctatcg ggggaagccc 120  
 51 tcgctgcgcc ccagggggag ctggcgatgt gaccagggtta agtgataacc atcaccttgc 180  
 52 caatggggttt gcgaacttta ccgtgacgct acccccgctt ttgtttgatc acaccagctc 240  
 53 gaaggtgtgc gcttttccga agatgcacgt gaagtggcaa atccttgcca cccgaggttt 300  
 54 tcccagtaca aacgtactag tgatgaggat cacggggaac attgtggaga ttgcactttg 360  
 55 caatatttgc aaaaggggtg actacccccg cgcaaaactt aaaaacccaa atccgttgac 420  
 56 ggacccatgc ccgatgaagc aatgtgtgaa gcacgccacc ggaacacagg ttgtggatca 480

## RAW SEQUENCE LISTING

DATE: 10/18/2004

PATENT APPLICATION: US/10/089,057A

TIME: 09:38:04

Input Set : A:\PTO.FG.txt

Output Set: N:\CRF4\10182004\J089057A.raw

```

57 ctcacccatga tgtgggggat tcgcatcaca cagtgtgcag ggcggcacct ctaccgaatg 540
58 cgccttacag cagcaccaag aagaagtgcac tcttag atg tca aac gtt gga acg 594
59                                     Met Ser Asn Val Gly Thr
60                                     1           5
61 cca cgt acc gca cag gaa atc cag cag gat tgg gac acc aac cca cgc 642
62 Pro Arg Thr Ala Gln Glu Ile Gln Gln Asp Trp Asp Thr Asn Pro Arg
63         10           15           20
64 tgg aac gga atc acc cgc gac tac acc gct gag cag gta gct gag ctc 690
65 Trp Asn Gly Ile Thr Arg Asp Tyr Thr Ala Glu Gln Val Ala Glu Leu
66         25           30           35
67 cag ggc agc gtc gtc gag gag cac acc ctc gca aag cgc ggc gcc gag 738
68 Gln Gly Ser Val Val Glu Glu His Thr Leu Ala Lys Arg Gly Ala Glu
69         40           45           50
70 atc ctg tgg gat gca gtt tcc gca gag ggc gac gac tac atc aac gca 786
71 Ile Leu Trp Asp Ala Val Ser Ala Glu Gly Asp Asp Tyr Ile Asn Ala
72 55           60           65           70
73 ctg ggc gcc ctt acc ggt aac cag gct gtc cag cag gtc cgt gcc ggc 834
74 Leu Gly Ala Leu Thr Gly Asn Gln Ala Val Gln Gln Val Arg Ala Gly
75         75           80           85
76 ctg aag gct gtc tac ctc tcc ggc tgg cag gtc gca ggt gac gcc aac 882
77 Leu Lys Ala Val Tyr Leu Ser Gly Trp Gln Val Ala Gly Asp Ala Asn
78         90           95           100
79 ctc gcc ggt cac acc tac ccc gac cag tcc ctg tac ccg gcg aac tcc 930
80 Leu Ala Gly His Thr Tyr Pro Asp Gln Ser Leu Tyr Pro Ala Asn Ser
81 105           110           115
82 gtc ccg aac gtt gtc cgt cgc atc aac aac gca ctg ctg cgc gcc gat 978
83 Val Pro Asn Val Val Arg Arg Ile Asn Asn Ala Leu Leu Arg Ala Asp
84 120           125           130
85 gag atc gca cgc gtc gag ggt gac acc tcc gtc gac aac tgg ctc gtc 1026
86 Glu Ile Ala Arg Val Glu Gly Asp Thr Ser Val Asp Asn Trp Leu Val
87 135           140           145           150
88 ccg atc gtc gcc gac ggc gag gcc ggc ttc ggt ggc gcc ctc aac gtc 1074
89 Pro Ile Val Ala Asp Gly Glu Ala Gly Phe Gly Gly Ala Leu Asn Val
90 155           160           165
91 tac gag ctc cag aag ggc atg atc acc gct ggt gcc gca ggc acc cac 1122
92 Tyr Glu Leu Gln Lys Gly Met Ile Thr Ala Gly Ala Ala Gly Thr His
93 170           175           180
94 tgg gag gat cag ctc gct tcc gag aag aag tgt ggc cac ctc ggt ggc 1170
95 Trp Glu Asp Gln Leu Ala Ser Glu Lys Lys Cys Gly His Leu Gly Gly
96 185           190           195
97 aag gtc ctc atc ccg acc cag cag cac atc cgc acc ctg aac tcc gcc 1218
98 Lys Val Leu Ile Pro Thr Gln Gln His Ile Arg Thr Leu Asn Ser Ala
99 200           205           210
100 cgc ctg gca gct gac gtg gcc aac acc ccg acc gtc gtc atc gcc cgc 1266
101 Arg Leu Ala Ala Asp Val Ala Asn Thr Pro Thr Val Val Ile Ala Arg
102 215           220           225           230
103 acc gac gca gag gcc gcc acc ctg atc acc tct gat gtt gat gag cgc 1314
104 Thr Asp Ala Glu Ala Ala Thr Leu Ile Thr Ser Asp Val Asp Glu Arg
105 235           240           245

```

## RAW SEQUENCE LISTING

DATE: 10/18/2004

PATENT APPLICATION: US/10/089,057A

TIME: 09:38:04

Input Set : A:\PTO.FG.txt

Output Set: N:\CRF4\10182004\J089057A.raw

```

106 gac cgc cca ttc atc acc ggc gag cgc acc gcc gag ggc tac tac cac 1362
107 Asp Arg Pro Phe Ile Thr Gly Glu Arg Thr Ala Glu Gly Tyr Tyr His
108          250          255          260
109 gtc aag ccg ggt ctc gag ccc tgc atc gca cgt gcg aag tcc tac gct 1410
110 Val Lys Pro Gly Leu Glu Pro Cys Ile Ala Arg Ala Lys Ser Tyr Ala
111          265          270          275
112 ccc tac gca gac atg atc tgg atg gag acc ggc acc cct gac ctc gag 1458
113 Pro Tyr Ala Asp Met Ile Trp Met Glu Thr Gly Thr Pro Asp Leu Glu
114          280          285          290
115 ctg gcc aag aag ttc gcc gag ggc gtc cgc agc gag ttc ccg gac cag 1506
116 Leu Ala Lys Lys Phe Ala Glu Gly Val Arg Ser Glu Phe Pro Asp Gln
117 295          300          305          310
118 ctg ctg tcc tac aac tgc tcc ccg tcc ttc aac tgg tct gca cac ctc 1554
119 Leu Leu Ser Tyr Asn Cys Ser Pro Ser Phe Asn Trp Ser Ala His Leu
120          315          320          325
121 gag gcc gac gag atc gct aag ttc cag aag gaa ctg ggt gcc atg ggc 1602
122 Glu Ala Asp Glu Ile Ala Lys Phe Gln Lys Glu Leu Gly Ala Met Gly
123          330          335          340
124 ttc aag ttc cag ttc atc acc ctg gct ggc ttc cac tcc ctc aac tac 1650
125 Phe Lys Phe Gln Phe Ile Thr Leu Ala Gly Phe His Ser Leu Asn Tyr
126          345          350          355
127 ggt atg ttc gac ctg gct tac ggc tac gcc cgt gaa ggc atg ccc gcc 1698
128 Gly Met Phe Asp Leu Ala Tyr Gly Tyr Ala Arg Glu Gly Met Pro Ala
129          360          365          370
130 ttc gtc gac ctg cag aac cgt gag ttc aag gca gct gag gag cgc ggc 1746
131 Phe Val Asp Leu Gln Asn Arg Glu Phe Lys Ala Ala Glu Glu Arg Gly
132 375          380          385          390
133 ttc acc gcc gtc aag cac cag cgt gag gtc ggc gcc ggc tac ttc gac 1794
134 Phe Thr Ala Val Lys His Gln Arg Glu Val Gly Ala Gly Tyr Phe Asp
135          395          400          405
136 acc atc gcc acc acc gtt gac ccg aac tcc tcc acc acc gcg ctg aag 1842
137 Thr Ile Ala Thr Thr Val Asp Pro Asn Ser Ser Thr Thr Ala Leu Lys
138          410          415          420
139 ggt tcc acc gag gaa tgc cag ttc cac taggaaccac ctgatgcggt 1889
140 Gly Ser Thr Glu Glu Cys Gln Phe His
141          425          430
142 gccgtatggc ctgacggcac cgcctctccc tttgcactcc agtactcctt tgtgcacatc 1949
143 ggccatctcc acaccgcgcg ccccgccacc t 1980
145 <210> SEQ ID NO: 2
146 <211> LENGTH: 431
147 <212> TYPE: PRT
148 <213> ORGANISM: Corynebacterium thermoaminogenes
150 <400> SEQUENCE: 2
151 Met Ser Asn Val Gly Thr Pro Arg Thr Ala Gln Glu Ile Gln Gln Asp
152 1          5          10          15
153 Trp Asp Thr Asn Pro Arg Trp Asn Gly Ile Thr Arg Asp Tyr Thr Ala
154          20          25          30
155 Glu Gln Val Ala Glu Leu Gln Gly Ser Val Val Glu Glu His Thr Leu
156          35          40          45

```

## RAW SEQUENCE LISTING

DATE: 10/18/2004

PATENT APPLICATION: US/10/089,057A

TIME: 09:38:04

Input Set : A:\PTO.FG.txt

Output Set: N:\CRF4\10182004\J089057A.raw

```

157 Ala Lys Arg Gly Ala Glu Ile Leu Trp Asp Ala Val Ser Ala Glu Gly
158      50                      55                      60
159 Asp Asp Tyr Ile Asn Ala Leu Gly Ala Leu Thr Gly Asn Gln Ala Val
160 65                      70                      75                      80
161 Gln Gln Val Arg Ala Gly Leu Lys Ala Val Tyr Leu Ser Gly Trp Gln
162      85                      90                      95
163 Val Ala Gly Asp Ala Asn Leu Ala Gly His Thr Tyr Pro Asp Gln Ser
164      100                     105                     110
165 Leu Tyr Pro Ala Asn Ser Val Pro Asn Val Val Arg Arg Ile Asn Asn
166      115                     120                     125
167 Ala Leu Leu Arg Ala Asp Glu Ile Ala Arg Val Glu Gly Asp Thr Ser
168      130                     135                     140
169 Val Asp Asn Trp Leu Val Pro Ile Val Ala Asp Gly Glu Ala Gly Phe
170 145                     150                     155                     160
171 Gly Gly Ala Leu Asn Val Tyr Glu Leu Gln Lys Gly Met Ile Thr Ala
172      165                     170                     175
173 Gly Ala Ala Gly Thr His Trp Glu Asp Gln Leu Ala Ser Glu Lys Lys
174      180                     185                     190
175 Cys Gly His Leu Gly Gly Lys Val Leu Ile Pro Thr Gln Gln His Ile
176      195                     200                     205
177 Arg Thr Leu Asn Ser Ala Arg Leu Ala Ala Asp Val Ala Asn Thr Pro
178      210                     215                     220
179 Thr Val Val Ile Ala Arg Thr Asp Ala Glu Ala Ala Thr Leu Ile Thr
180 225                     230                     235                     240
181 Ser Asp Val Asp Glu Arg Asp Arg Pro Phe Ile Thr Gly Glu Arg Thr
182      245                     250                     255
183 Ala Glu Gly Tyr Tyr His Val Lys Pro Gly Leu Glu Pro Cys Ile Ala
184      260                     265                     270
185 Arg Ala Lys Ser Tyr Ala Pro Tyr Ala Asp Met Ile Trp Met Glu Thr
186      275                     280                     285
187 Gly Thr Pro Asp Leu Glu Leu Ala Lys Lys Phe Ala Glu Gly Val Arg
188      290                     295                     300
189 Ser Glu Phe Pro Asp Gln Leu Leu Ser Tyr Asn Cys Ser Pro Ser Phe
190 305                     310                     315                     320
191 Asn Trp Ser Ala His Leu Glu Ala Asp Glu Ile Ala Lys Phe Gln Lys
192      325                     330                     335
193 Glu Leu Gly Ala Met Gly Phe Lys Phe Gln Phe Ile Thr Leu Ala Gly
194      340                     345                     350
195 Phe His Ser Leu Asn Tyr Gly Met Phe Asp Leu Ala Tyr Gly Tyr Ala
196      355                     360                     365
197 Arg Glu Gly Met Pro Ala Phe Val Asp Leu Gln Asn Arg Glu Phe Lys
198      370                     375                     380
199 Ala Ala Glu Glu Arg Gly Phe Thr Ala Val Lys His Gln Arg Glu Val
200 385                     390                     395                     400
201 Gly Ala Gly Tyr Phe Asp Thr Ile Ala Thr Thr Val Asp Pro Asn Ser
202      405                     410                     415
203 Ser Thr Thr Ala Leu Lys Gly Ser Thr Glu Glu Cys Gln Phe His
204      420                     425                     430
206 <210> SEQ ID NO: 3

```

## RAW SEQUENCE LISTING

DATE: 10/18/2004

PATENT APPLICATION: US/10/089,057A

TIME: 09:38:04

Input Set : A:\PTO.FG.txt

Output Set: N:\CRF4\10182004\J089057A.raw

```

207 <211> LENGTH: 2381
208 <212> TYPE: DNA
209 <213> ORGANISM: Corynebacterium thermoaminogenes
211 <220> FEATURE:
212 <221> NAME/KEY: CDS
213 <222> LOCATION: (577)..(2349)
215 <400> SEQUENCE: 3
216 agcaggccgt gttgccgaac ggcaacttcc gcagccgcaa ggagatcgag gaggtgtact 60
217 cgcacctcaa ccttgccgag gacaccgtgg tgtactgccg cgtgggtgac cgcgcggccc 120
218 acacctgggt cgtgttgaag tacctgctgg ggtttgaaaa cgtccgcaac tatgacgggt 180
219 cctgggtccga gtggggcaac atggtgcgca tgcccatcgt ccagggtgat gagccgggct 240
220 cactctagtc accccggggt cacctccctg gtcacccccg taccctcccg ggtacacccc 300
221 ggggacgggg tgtgacctgg atctccctg catgtggaca ccgggaaact ttgcctggga 360
222 aatgaccatc cagtaccgta atgcgggtat gttaacgcgg tcacagggtg caccagaatc 420
223 cggatcgtct aacccccctt gcgggattcg ctaaaagatc accgagttag tgtgcaagaa 480
224 taatgctgat cgcaggggca ctgtcatacg ctgtcatgca gtcaatgaac agtgcggtgc 540
225 tctgtcgtga agaaaatcaa aaccaggagg gtttta gtg tca gtc gag acc agg 594
226                                     Val Ser Val Glu Thr Arg
227                                     1 5
228 aag atc acc aag gta ctt gtc gcc aac cgt ggt gaa atc gca atc cgt 642
229 Lys Ile Thr Lys Val Leu Val Ala Asn Arg Gly Glu Ile Ala Ile Arg
230 10 15 20
231 gtt ttc cgc gca gca cgg gat gaa ggc atc gcc tct gtc gcc gtc tac 690
232 Val Phe Arg Ala Ala Arg Asp Glu Gly Ile Ala Ser Val Ala Val Tyr
233 25 30 35
234 gcg gag ccg gac gca gat gcc cct ttc gtc gag tat gcc gat gag gcc 738
235 Ala Glu Pro Asp Ala Asp Ala Pro Phe Val Glu Tyr Ala Asp Glu Ala
236 40 45 50
237 ttc gca ctc ggt ggc cag act tcc gca gag tcc tac ctc gtc att gac 786
238 Phe Ala Leu Gly Gly Gln Thr Ser Ala Glu Ser Tyr Leu Val Ile Asp
239 55 60 65 70
240 aag atc att gac gca gca cgc aag tcc ggt gca gac gct gtc cac ccc 834
241 Lys Ile Ile Asp Ala Ala Arg Lys Ser Gly Ala Asp Ala Val His Pro
242 75 80 85
243 ggc tac ggc ttc ctc gcc gag aac gcc gat ttc gct gaa gct gtc atc 882
244 Gly Tyr Gly Phe Leu Ala Glu Asn Ala Asp Phe Ala Glu Ala Val Ile
245 90 95 100
246 aac gag ggc ctg atc tgg atc gga cca tcc cct gag tcc atc cgt tcc 930
247 Asn Glu Gly Leu Ile Trp Ile Gly Pro Ser Pro Glu Ser Ile Arg Ser
248 105 110 115
249 ctc ggt gac aag gtc acc gca cgc cac atc gcc aac aac gcc aac gca 978
250 Leu Gly Asp Lys Val Thr Ala Arg His Ile Ala Asn Asn Ala Asn Ala
251 120 125 130
252 ccg atg gca ccg ggc acc aag gag cct gtc aag gac gcc gct gag gtt 1026
253 Pro Met Ala Pro Gly Thr Lys Glu Pro Val Lys Asp Ala Ala Glu Val
254 135 140 145 150
255 gtc gcc ttc gcc gag gag ttc ggt ctc ccc atc gcc atc aag gct gcc 1074
256 Val Ala Phe Ala Glu Glu Phe Gly Leu Pro Ile Ala Ile Lys Ala Ala
257 155 160 165

```

RAW SEQUENCE LISTING ERROR SUMMARY      DATE: 10/18/2004  
PATENT APPLICATION: US/10/089,057A      TIME: 09:38:05

Input Set : A:\PTO.FG.txt  
Output Set: N:\CRF4\10182004\J089057A.raw

**Please Note:**

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:45; N Pos. 3  
Seq#:46; N Pos. 18  
Seq#:55; N Pos. 3,6,9  
Seq#:56; N Pos. 3,9,18  
Seq#:83; N Pos. 9

VERIFICATION SUMMARY

DATE: 10/18/2004

PATENT APPLICATION: US/10/089,057A

TIME: 09:38:05

Input Set : A:\PTO.FG.txt

Output Set: N:\CRF4\10182004\J089057A.raw

18 M:283 W: Missing Blank Line separator, <130> field identifier  
3855 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:45 after pos.:0  
3871 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:46 after pos.:0  
3985 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:55 after pos.:0  
4011 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:56 after pos.:0  
4606 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:83 after pos.:0